

10042865 Results

SEQ ID NO: 28

Result No.	Score	Query Match	Length	DB	ID	Description
1	2824	100.0	533	5	ABB98414	Abb98414 Human NOV
2	1794	63.5	529	5	AAO22794	Aao22794 Protein o
3	1786	63.2	529	5	AAE28617	Aae28617 Human UGT
4	1786	63.2	529	6	ADA11076	Ada11076 Human cDN
5	1784	63.2	533	4	ABG05523	Abg05523 Novel hum
6	1771.5	62.7	528	3	AAAY78933	Aay78933 Human UDP
7	1770	62.7	524	3	AAAY78934	Aay78934 Human UDP
8	1753.5	62.1	540	4	ABG05525	Abg05525 Novel hum
9	1751.5	62.0	530	7	ADC39065	Adc39065 Novel hum
10	1741.5	61.7	530	2	AAW47126	Aaw47126 Uridine d
11	1741.5	61.7	530	7	ADE58009	Ade58009 Human Pro
12	1732.5	61.3	530	3	AAAY78935	Aay78935 Human UDP
13	1732.5	61.3	530	6	ABJ19806	Abj19806 Androgen-
14	1716	60.8	529	4	AAE02188	Aae02188 Human bre
15	1709	60.5	532	4	ABG05524	Abg05524 Novel hum

RESULT 6

AAAY78933

ID AAY78933 standard; protein; 528 AA.

XX

AC AAY78933;

XX

DT 05-JUN-2000 (first entry)

XX

DE Human UDP-glucuronosyltransferase 2B4 amino acid sequence.

XX

KW UDP-glucuronosyltransferase 2B4; UGT2B4; polymorphism; metabolism; SNPs;
KW drug interaction; detect; human; single nucleotide polymorphism.

XX

OS Homo sapiens.

XX

PN WO200006776-A1.

XX

PD 10-FEB-2000.

XX

PF 22-JUL-1999; 99WO-US016675.

XX

PR 28-JUL-1998; 98US-0094391P.

XX

PA (AXYS-) AXYS PHARM INC.

XX

PI Galvin M, Miller A, Penny L, Riedy M;

XX

DR WPI; 2000-195321/17.

DR N-PSDB; AAZ95199.

XX

PT Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
PT genotyping individuals to predict rate of metabolism of substrates and
PT for identifying potential drug interactions.

XX

PS Disclosure; Page 36-37; 72pp; English.

XX

CC This sequence represents the human UDP-glucuronosyltransferase 2B4
 CC (UGT2B4) amino acid sequence. UDP-glucuronosyltransferase (UGTs) are a
 CC family of enzymes that catalyse the glucuronic acid conjugation of a wide
 CC range of endogenous and exogenous substrates. The UGT2B gene subfamily
 CC encode steroid metabolizing isoforms in the liver. Alteration of the
 CC expression or function of UGTs may effect drug metabolism. The invention
 CC relates to non-chromosomal nucleic acid molecules, which comprise human
 CC UGT2B sequence polymorphisms (see AAZ95051-Z95110). Probes which detect

CC the UGT2B locus polymorphisms can be used to detect altered UGT2B
 CC metabolism of a substrate in an individual. The nucleic acid molecules
 CC comprising a human UGT2B sequence polymorphism can be used in screening
 CC assays for genotyping individuals, also to predict their rate of
 CC metabolism of UGT2B substrate, potential drug-drug interactions and
 CC adverse side effects. The polymorphisms can be used as single nucleotide
 CC polymorphisms (SNPs) for detecting genetic linkage related to phenotypic
 CC variation in activity or expression of UGT2B protein. The polymorphism
 CC containing nucleic acid molecules may also be used for generating
 CC genetically modified non-human animals and for obtaining site specific
 CC gene modification in cell lines

XX

SQ Sequence 528 AA;

Query Match 62.7%; Score 1771.5; DB 3; Length 528;
 Best Local Similarity 66.2%; Pred. No. 1.3e-173;
 Matches 355; Conservative 54; Mismatches 116; Indels 11; Gaps 9;

Qy 1 MAMKWTSVLLLLIQLSYSSSGCGNVLPWPMYSPWMNIKTILDKLMQISHEVTVLTLSA 60
 | : ||||| ||||| | : || | : ||||| : | : ||||| ||
 Db 1 MSMKWTSAALLIQLSCYFSSSGCGKLVLPWPTFESHWMNIKTILDELVQRGHEVTVLASSA 60

Qy 61 SILVDPNITSVTKFVYISIVIKDDFAGFFFTQQITKWIHDLPKHIFWFKCVPFKNILWE 120
 || | || | ||||| : | : | : | : | : || | : | : |
 Db 61 SISFDPNPSTLKFVYPVSLTKTEFED-IIKQLVKRWA-ELPKDTFWSYFSQVQEIMWT 118

Qy 121 YSGYTEKFFKDVVLNKKLMTNLQESRSDVVHANAIGPFGELLAELLKISFVYSLHFSPGY 180
 : : || | : ||||| ||||| || | : : ||||| ||||| ||||| |||||
 Db 119 FNDILRKFKCDIVSNKKLKKLQESRFDVVLADAVFPFGELLAELLKIPFVYSLRFSFGY 178

Qy 181 TFEKYSGGFLLPPSYGAVILSELGSGMTFMETVRNIYVYFDFWFQTFDMKKGDQFYSE 240
 || : |||| | |||| | : |||| |||| : | : |||| ||||| ||||| |||||
 Db 179 AIEKHSGGLFPFSYVPVMSSELSDQMTFIERVKNNIYVLYFEFQIFDMKKWDQFYSE 238

Qy 241 VLGKSCFLSEIMGKAEMWLIRNYWYLEFPRPLLPNFEFVRLYCKPVNPLPKEKMEEFAQ 300
 ||| : || | || : ||||| : || ||||| || | : |||| ||||| |||||
 Db 239 VLGRPTTLSETMAKADIWLIRNYWDFQFPHPLLPNVEFVGGLHCKPAKPLPKE-MEEFVQ 297

Qy 301 SSDEDG-VVFSLESVQNLTTEKADLITSALAQIPQKVM-KF-GRKPNTLRNTQWHRWI 357
 || | : |||| | | | : || : || ||||| : || : || : || : || : || : ||
 Db 298 SSGENGVVVFSGLGSMVSNTESEERANVIASALAKIPQKVLWRFDGKNKPDITLGLNTRYKWI 357

Qy 358 PQNECLILDHPQTKAFITYGGTNSIYEMIYRGVPSMGIFLADQHDNIAHMKAKGAAVIL 417
 ||| : : || : || : ||||| | || | || | : | : ||||| ||||| ||||| |||||
 Db 358 PQND--LLGHPKTRAFITHGGANGIYEAHYHGIPMVGVPFLFADQPDNIAHMKAKGAAVSL 415

Qy 418 DLSTKSTDLDISVFSVLSFLSFRYKESVMKLSRIQHDQPVKPLDRAVFWIEFVMRHKGA 477
 | | ||||| : : || : ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 416 DFHTMSSTDLL--NALKTVINDPLYKENAMKLSRIHHDQPVKPLDRAVFWIEFVMRHKGA 473

Qy 478 KHLRVAARDLTWFQYHSLDVIGFLLACVATVTFTIITKCCFLCFWKFTRKVKKEKRD 533
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 474 KHLRVAADLTWFQYHSLDVTGFLACVATVIFTITK-CLFCVWKFVRTGKKGKRD 528

Issued

SUMMARIES

Result	% Query					
No.	Score	Match	Length	DB	ID	Description
1	1771.5	62.7	528	4	US-09-356-806-8	Sequence 8, Appli
2	1770	62.7	524	4	US-09-356-806-40	Sequence 40, Appl
3	1741.5	61.7	530	3	US-09-180-852-2	Sequence 2, Appli
4	1732.5	61.3	530	4	US-09-356-806-113	Sequence 113, App
5	1375.5	48.7	454	4	US-09-813-918-2	Sequence 2, Appli
6	998	35.3	288	4	US-09-813-918-3	Sequence 3, Appli
7	779	27.6	531	5	PCT-US92-00282-6	Sequence 6, Appli
8	777.5	27.5	534	5	PCT-US92-00282-4	Sequence 4, Appli
9	777	27.5	533	5	PCT-US92-00282-3	Sequence 3, Appli

10	737.5	26.1	531	5	PCT-US92-00282-5	Sequence 5, Appli
11	689.5	24.4	529	5	PCT-US92-00282-7	Sequence 7, Appli
12	568.5	20.1	245	4	US-09-305-856B-18	Sequence 18, Appl
13	543	19.2	197	4	US-09-813-918-4	Sequence 4, Appli
14	311	11.0	98	5	PCT-US92-00282-26	Sequence 26, Appl
15	281	10.0	515	3	US-08-942-012B-32	Sequence 32, Appl

Result No.	Score	Query Match	Length	DB	ID	Description
1	1756	62.2	529	6	Q97951	Q97951 macaca fasc
2	1745	61.8	529	6	Q9GLD9	Q9gld9 macaca mula
3	1734	61.4	529	6	Q9TSL6	Q9tsl6 macaca fasc
4	1730	61.3	529	6	Q9GLE0	Q9gle0 macaca mula
5	1700.5	60.2	528	6	Q8WN97	Q8wn97 macaca fasc
6	1554.5	55.0	529	11	Q8R084	Q8r084 mus musculu
7	1554.5	55.0	532	11	Q8K154	Q8k154 mus musculu
8	1545.5	54.7	528	11	Q8VIF9	Q8vif9 cavia porce
9	1481	52.4	529	11	Q8VIF8	Q8vif8 cavia porce
10	1473	52.2	529	11	Q8BJL9	Q8bjl9 mus musculu
11	1465.5	51.9	530	11	Q8K169	Q8k169 mus musculu

Result No.	Score	Query Match	Length	DB	ID	Description
1	1786	63.2	529	1	UDB7_HUMAN	P16662 homo sapien
2	1771.5	62.7	528	1	UDB4_HUMAN	P06133 homo sapien
3	1741.5	61.7	530	1	UDBH_HUMAN	O75795 homo sapien
4	1741	61.7	529	1	UDB9_MACFA	O02663 macaca fasc
5	1732.5	61.3	530	1	UDBF_HUMAN	P54855 homo sapien
6	1723.5	61.0	528	1	UDBA_HUMAN	P36537 homo sapien
7	1721.5	61.0	528	1	UDBJ_MACFA	Q9xt55 macaca fasc
8	1716	60.8	529	1	UDBB_HUMAN	O75310 homo sapien
9	1683	59.6	529	1	UDBS_HUMAN	Q9by64 homo sapien
10	1677.5	59.4	530	1	UDBK_MACFA	O77649 macaca fasc
11	1563.5	55.4	523	1	UDBG_RABIT	O19103 oryctolagus
12	1553	55.0	531	1	UDBD_RABIT	P36512 oryctolagus
13	1545.5	54.7	529	1	UDB1_RAT	P09875 rattus norv
14	1540.5	54.6	530	1	UDBE_RABIT	P36513 oryctolagus
15	1508.5	53.4	530	1	UDBC_RAT	P36511 rattus norv
16	1461.5	51.8	530	1	UDB5_MOUSE	P17717 mus musculu
17	1445.5	51.2	530	1	UDB2_RAT	P08541 rattus norv
18	1437.5	50.9	530	1	UDB3_RAT	P08542 rattus norv
19	1413.5	50.1	530	1	UDB6_RAT	P19488 rattus norv

SEQ ID NO : 27

Result No.	Score	Query Match	Length	DB	ID	Description
1	1606	100.0	1606	6	AX675577	AX675577 Sequence
2	1606	100.0	1606	6	AX921811	AX921811 Sequence
3	973	60.6	1855	6	AX336329	AX336329 Sequence
4	973	60.6	1855	6	AX336696	AX336696 Sequence
5	973	60.6	1855	6	AX409473	AX409473 Sequence
6	973	60.6	1855	9	HUMUDPGTA	J05428 Human 3,4-c
7	966.6	60.2	1766	9	BC030974	BC030974 Homo sapi
8	966.6	60.2	1854	6	BD229166	BD229166 Genotype
9	966.6	60.2	1854	6	AR349418	AR349418 Sequence
10	962.4	59.9	1639	6	AX548042	AX548042 Sequence
11	958.6	59.7	2107	6	AR168316	AR168316 Sequence
12	958.6	59.7	2107	9	HSU59209	U59209 Homo sapien
13	953.8	59.4	1753	9	AF016310	AF016310 Macaca fa
14	950.6	59.2	1976	6	BD229238	BD229238 Genotype
15	950.6	59.2	1976	6	AR349490	AR349490 Sequence

RESULT 6
HUMUDPGTA
LOCUS HUMUDPGTA 1855 bp mRNA linear PRI 03-AUG-1993
DEFINITION Human 3,4-catechol estrogen UDP-glucuronosyltransferase mRNA, complete cds.
ACCESSION J05428
VERSION J05428.1 GI:340079
KEYWORDS 3,4-catechol estrogen UDP-glucuronosyltransferase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1855)
AUTHORS Ritter,J.K., Sheen,Y.Y. and Owens,I.S.
TITLE Cloning and expression of human liver UDP-glucuronosyltransferase in COS-1 cells. 3,4-catechol estrogens and estriol as primary substrates
JOURNAL J. Biol. Chem. 265 (14), 7900-7906 (1990)
MEDLINE 90243659
PUBMED 2159463
COMMENT Original source text: Human liver, cDNA to mRNA, clone 63-11. Draft entry and computer-readable sequence for [1] kindly submitted by I.S.Owens, 22-FEB-1990.
FEATURES
source Location/Qualifiers
1..1855
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
CDS
15..1604
/note="UDP-glucuronosyltransferase (EC 2.4.1.17)"
/codon_start=1
/protein_id="AAA36793.1"
/db_xref="GI:340080"
/translation="MSVKWTSVILLIQLSFCFSSGNCCKVLVWAAEYSHWMNIKTILD
ELIQRGHEVTVLASSASILFDPNNSSALKIEIYPTSLTKTELENFIMQQIKRWSDLPK
DTFWLYFSQVQEIMSIIFGDITRKFCCKDVVSNNKKFMKKVQESRFDVIFADAIIFPCSELL
AELFNIPFVYSLSPGYTFEKHSGGFIFPPSYVPVVMSELTDMQTFMERVKNMIIYVL
YFDWFEEIFDMKKWDQFYSEVLGRPTTLSETMGKADVWLIRNSWNFQFPHLLPNVDF
VGGLHCKPAKPLPKEMEDFVQSSGENGVVVFSLGSMVSNMTEERANVIASALAQIPQK
VLWRFDGNKPDTLGLNTRLYKWIQNDLLGHPKTRAFITGGANGIYEAIYHGIPMVG
IPLFADQPDNIAHMKARGAARVDFNTMSSTDLLNALKRVINDPSYKENVMKLSRIQH
DQPVKPLDRAVFWIEFVMRHKGAKHLRVAADLTWFQYHSLDVIGFLLVCVATVIFIV
TKCCLFCFWKFARKAKKGKND"

ORIGIN

Query Match 60.6%; Score 973; DB 9; Length 1855;
Best Local Similarity 78.3%; Pred. No. 1e-212;
Matches 1265; Conservative 0; Mismatches 320; Indels 30; Gaps 7;

Qy 1 ATGGCTATGAAATGGACTTCAGTCCTTCTGTTGATACAGCTGAGCTATTACTCTAGCTCT 60
||| || ||||| ||||| || || ||||| ||||| || || |||||
Db 15 ATGCTCTGTGAAATGGACTTCAGTAATTTTGCTAATAACAAGTGGCTTTTGCTTTAGCTCT 74
Qy 61 GGGAGTTGTGGAAATGTGCCGCTGTGGCCCATGGAATATAGTCCTTGGATGAATATAAAG 120
|||| ||||| ||||| ||||| || ||||| || ||||| ||||| |||||
Db 75 GGGAAATTGTGGAAAGGTGCTGGTGTGGGCAGCAGAATACAGCCATTGGATGAATATAAAG 134
Qy 121 ACAATCCTGGATAAACTTATGCAGATAAGTCATGAGGTGACTGTTCTAACATTGTGAGCT 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 ACAATCCTGGATGAGCTTATTGAGAGAGTCATGAGGTGACTGTACTGGCATCTTCAGCT 194
Qy 181 TCCATTCTTGTGTGATCCCAACATAACATCTGTTACTAAATTTGAGGTTTATTCTATATCT 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 TCCATTCTTTTGTGATCCCAACAACTCATCGCTCTTAAATTTGAAATTTATCCACATCT 254
Qy 241 GTAATTAAAGATGATTTTGCAGGTTTTCACACAACAGATTACTAAATGGATACAT 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 TTAATAAAAGTGGTT---GGAGAATTTCATCATGCAACAGATTAAGAGATGGTCA--- 308
Qy 301 GATCTTCCAAAACATATATTTTGGTTTAAATGTGTTCCCTTCAAGAATATTCTTTGGGAA 360
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 309 GACCTTCCAAAAGATACATTTTGGTTATATTTTTCACAAGTACAGGAAATCATGTCAATA 368
 Qy 361 TATTCTGGTTATACTGAGAAGTCTTTTAAAGATGTAGTTTGAACAAGAACTTATGACA 420
 Db 369 TTTGGTGACATAACTAGAAAGTCTGTAAAGATGTAGTTTCAAATAAGAAATTTATGAAA 428
 Qy 421 AACCTACAAGAATCAAGGTCTGATGTCGTTTCATGCAAATGCCATTGGTCCCTTTGGAGAG 480
 Db 429 AAAGTACAAGAGTCAAGATTTGACGTCATTTTGCAGATGCTATTTTCCCTGTAGTGAG 488
 Qy 481 CTGCTGGCTGAGCTATTAAAAATATCCTTTGTGTACAGTCTCCACTTCTCTCCTGGCTAC 540
 Db 489 CTGCTGGCTGAGCTATTAAACATACCTTTGTGTACAGTCTCAGCTTCTCTCCTGGCTAC 548
 Qy 541 ACATTTGAGAAATACAGTGGAGGATTCTACTTCCACCTTCTATGGAGCTGTTATTCTG 600
 Db 549 ACTTTTGAAAGCATAGTGGAGGATTATTTTCCCTCCTTCTACGTACCTGTTGTTATG 608
 Qy 601 TCAGAATTAAGTGGTTCGATGACATTTCATGGAGACAGTAAGAAATATTATATATGTGTTT 660
 Db 609 TCAGAATTAAGTGGTTCGATGACATTTCATGGAGAGGTAAGAAATATGATCTATGTGCTT 668
 Qy 661 TATTTTGACTTTTGGTTCCAAACATTTGATATGAAGAAGGAGACCAGTTTACAGTGAA 720
 Db 669 TACTTTGACTTTTGGTTTCGAAATATTGACATGAAGAAGTGGGATCAGTTTATAGTGAA 728
 Qy 721 GTTCTAGGTAAGTCATGTTTTTTATCTGAGATAATGGGAAAGCTGAAATGTGGCTCATT 780
 Db 729 GTTCTAGGAAGACCCACTACGTTATCTGAGACAATGGGAAAGCTGACGTATGGCTTATT 788
 Qy 781 CGAAACTACTGGTATTTGGAATTTCTCGCCCACTTTACCTAATTTGAATTTGTTGTA 840
 Db 789 CGAAACTCCTGGAATTTTCAGTTTCTCATCCACTTTACCAAATGTTGATTTTGTGGA 848
 Qy 841 AGACTCTACTGCAAACCTGTCAACCCCTGCCTAAGGAGAAAATGGAAGAATTTGCCAG 900
 Db 849 GGACTCCACTGCAAACCTGCCAAACCCCTGCCTAAGG---AAATGGAAGACTTTGTACAG 905
 Qy 901 AGCTCTGATGAAGACGGTGTT---GTGTTTTCTCTGGAGTCAGCTGTGCAAAACCTTACA 957
 Db 906 AGCTCTGGAGAAAATGGTGTGTGGTGTCTCTGGGGTCAATGGTCAGTAACATGACA 965
 Qy 958 GAAGAAAAGCTGATCTTATCACTTCGGCCCTGGCTCAGATTCCACAAAAGTCATGAAG 1017
 Db 966 GAAGAAAGGCCAACGTAATTGCATCAGCCCTGGCCAGATCCCACAAAAGTTCTGTGG 1025
 Qy 1018 -----TTCGAAGGAAACCAAATACCTTAAGATCCAATACTCAGTGGCATAGGTGGATC 1071
 Db 1026 AGATTGTATGGGAATAAACAGATACCTTAGGTCTCAATACTCGGCTGTATAAGTGGATA 1085
 Qy 1072 CCACAGAATGAATGTCTTATCCTAGATCATCCCCAAACCAAAGCCTTTATAACTTATGGT 1131
 Db 1086 CCCCAGAATGA-----CCTTCTAGGTTCATCAAAGACCAGAGCTTTTATAACTCATGGT 1139
 Qy 1132 GGAACAAATAGCATCTATGAGATGATCTACCGTGGAGTCCCTTCCATGGGCATTCTTTG 1191
 Db 1140 GGAGCCAATGGCATCTACGAGGCAATCTACCATGGGATCCCTATGGTGGGATTCATTG 1199
 Qy 1192 TTTGCGGACCAACATGATAACATTGCTCACATGAAGGCCAAGGGAGCAGCTGTTATATTG 1251
 Db 1200 TTTGCCGATCAACCTGATAACATTGCTCACATGAAGGCCAGGGAGCAGCTGTTAGAGTG 1259
 Qy 1252 GACTTGAGCACAAAGTCAAGTACAGATTGCTCGATATATCTGTGTTCTGTTATTTT 1311
 Db 1260 GACTTCAACACAATGTCGAGTACAGACTTGCTGAATGCATTGAAGAGAGTAATTAATGAT 1319
 Qy 1312 TTATCCTTCAGATATAAAGAGAGTGTTATGAAATTATCAAGAATTCAACATGATCAACCA 1371
 Db 1320 CCTTC-----ATATAAAGAGAATGTTATGAAATTATCAAGAATTCAACATGATCAACCA 1373
 Qy 1372 GTGAAGCCCTGGATCGAGCAGTCTTCTGGATTGAATTGTGTCATGCGCCACAAAGGAGCC 1431

```

Db      1374  GTGAAGCCCCTGGATCGAGCAGTCTTCTGGATTGAATTTGTCATGCGCCACAAAGGAGCT 1433
Qy      1432  AAACACCTTCGAGTTGCAGCCCGTGACCTCACCTGGTTCCAGTACCACTCTTTGGATGTG 1491
Db      1434  AAACACCTTCGGGTTGCAGCCCAGACCTCACCTGGTTCCAGTACCACTCTTTGGATGTG 1493
Qy      1492  ATTGGGTTTCTGCTGGCCTGTGTGGCAACTGTGACATTTATCATCACAAAGTGTGTCTG 1551
Db      1494  ATTGGGTTCTCTGCTGGTCTGTGTGGCAACTGTGATATTTATCGTCACAAAATGTTGTCTG 1553
Qy      1552  TTTTGTCTTCTGGAAGTTTACTAGAAAAGTGAAGAAGGAAAAAAGGGATTAGTTAT 1606
Db      1554  TTTTGTCTTCTGGAAGTTTGTCTAGAAAAGCAAAGAAGGAAAAAATGATTAGTTAT 1608

```

RESULT 12

HSU59209

LOCUS HSU59209 2107 bp mRNA linear PRI 02-JUL-1998

DEFINITION Homo sapiens C19steroid specific UDP-glucuronosyltransferase mRNA, complete cds.

ACCESSION U59209

VERSION U59209.1 GI:3287472

KEYWORDS UGT2B17G.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2107)

AUTHORS Beaulieu,M., Levesque,E., Hum,D.W. and Belanger,A.

TITLE Isolation and characterization of a novel cDNA encoding a human UDP-glucuronosyltransferase active on C19 steroids

JOURNAL J. Biol. Chem. 271 (37), 22855-22862 (1996)

MEDLINE 96394358

PUBMED 8798464

REFERENCE 2 (bases 1 to 2107)

AUTHORS Hum,D.W., Belanger,A., Beaulieu,M. and Levesque,E.

TITLE Direct Submission

JOURNAL Submitted (23-MAY-1996) Laboratory of Molecular Endocrinology, Centre Hospitalier de l'universite Laval, 2705 Boul. Laurier, Ste-Foy, Quebec G1V 4G2, Canada

FEATURES Location/Qualifiers

source 1..2107

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_type="LNCaP"

/tissue_type="prostate"

CDS 52..1644

/note="UGT2B17"

/codon_start=1

/product="C19steroid specific UDP-glucuronosyltransferase"

/protein_id="AAC25491.1"

/db_xref="GI:3287473"

/translation="MSLKWMSVFLMLQLSCYFSSGSCGKVLVWPTEYSHWINMKTILE
ELVQRGHEVIVLTSSASILVNASKSSAIKLEVYPTSLTKNDLEDFMCKMFDRTYSIS
LAELLNIPFLYSLRFSVGTYTEKNGGGFLFPSPYVPVVMSELSQDMIFMERIKNMIYM
LYFDFWFQAYDLKKWDQFYSEVLGRPTTLFETMGKAEMWLIRTYWDFEFPFLPNVD
FVGGLHCKPAKPLPKEMEETFVQSSGENGIVVFSLSGSMISNMSEESANMIASALAIPQ
KVLWRFDGKKPNTLGSNTRLYKWLPQNDLLGHPKTKAFITHGGTNGIYEAIYHGIPMV
GIPLFADQHDNIAHMKAKGAALSVDIRTMSRDLLNALKSVINDPIYKENIMKLSRIH
HDQPVKPLDRAVFWIEFVMRHKGAKHLRVAHNLTWQYHSLDVIAFLACVATMIFM
ITKCLLFCFRKLAKTGKKKKRD"

ORIGIN

Query Match 59.7%; Score 958.6; DB 9; Length 2107;

Best Local Similarity 77.3%; Pred. No. 2.1e-209;

Matches 1249; Conservative 0; Mismatches 339; Indels 27; Gaps 6;

```

Qy      1  ATGGCTATGAAATGGACTTCAGTCCTTCTGTTGATACAGCTGAGCTATTACTCTAGCTCT 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      52  ATGTCTCTGAAATGGATGTGCTAGTCCTTCTGCTGATGCAGCTCAGTTGTTACTTTAGCTCT 111

```

Qy	61	GGGAGTTGTGGAAATGTGCCCGCTGTGGCCCATGGAATATAGTCCCTGGATGAATATAAAG	120
Db	112	GGGAGTTGTGGAAAGGTGCTGGTGTGGCCACAGAATACAGCCATTGGATAAAATATGAAG	171
Qy	121	ACAATCCTGGATAAACTTATGCAGATAAGTCATGAGGTGACTGTTCTAACATTGTGAGCT	180
Db	172	ACAATCCTGGAAGAGCTTGTTCAGAGGGGTCATGAGGTGATTGTGTTGACATCTTCGGCT	231
Qy	181	TCCATTCTTGTTGATCCCAACATAACATCTGTACTAAATTGAGGTTTATTCTATATCT	240
Db	232	TCTATTCTTGTCATAGCCAGTAAATCATCTGCTATTAAATTAGAAGTTTATCCTACATCT	291
Qy	241	GTAATTAAAGATGATTTTGCAGGGTTTTTTTTTCACACAACAGATTACTAAATGGATACAT	300
Db	292	TTAACTAAAAATGATTGGAAGATTTTTTTATGA---AAATGTTTCGATAGATGGACATAT	348
Qy	301	GATCTTCCAAAACATATATTTTTGGTTTAAATGTGTTCCCTTCAAGAAATATCTTTGGGAA	360
Db	349	AGTATTTCAAAAATACATTTTGGTCATATTTTTCACAACACACAAGAAATGTGTTGGGAA	408
Qy	361	TATTCTGGTTTATACTGAGAAGTCTTTAAAGATGTAGTTTTGAACAAGAACTTATGACA	420
Db	409	TATTCTGACTATAATATAAGCTCTGTGAAGATGCAGTTTTGAACAAGAACTTATGAGA	468
Qy	421	AACCTACAAGAAATCAAGGTCTGATGTCGTTTCATGCAAATGCCATTGGTCCCTTTGGAGAG	480
Db	469	AACTACAAGAGTCAAAATTTGATGTCCTTCTGGCAGATGCCGTTAATCCCTGTGGTGAG	528
Qy	481	CTGCTGGCTGAGCTATTAATAATATCCTTTGTGTACAGTCTCCACTTCTCTCCTGGCTAC	540
Db	529	CTGCTGGCTGAACACTACTTAACATACCCTTTCTGTACAGTCTCCGCTTCTCTGTGGCTAC	588
Qy	541	ACATTTGAGAAATACAGTGGAGGATTTCTACTTCCACCTTCCTATGGAGCTGTTATTCTG	600
Db	589	ACAGTTGAGAAGAAATGGTGGAGGATTTCTGTTCCCTCCTTCCTATGTACCTGTGTTATG	648
Qy	601	TCAGAATTAAGTGGTTCGATGACATTTCATGGAGACAGTAAGAAATATTATATATGTGTTT	660
Db	649	TCAGAATTAAGTGATCAAATGATTTTCATGGAGAGGATAAAAAATATGATATATATGCTT	708
Qy	661	TATTTTGACTTTTGGTTCCAAACATTTGATATGAAGAAGGAGACCAGTTTTACAGTGAA	720
Db	709	TATTTTGACTTTTGGTTTCAAGCATATGATCTGAAGAAGTGGGACCAGTTTTATAGTGAA	768
Qy	721	GTTCTAGGTAAAGTCATGTTTATCTGAGATAATGGGAAAAGCTGAAATGTGGCTCATT	780
Db	769	GTTCTAGGAAGACCCACTACATTATTGAGACAATGGGAAAAGCTGAAATGTGGCTCATT	828
Qy	781	CGAAACTACTGGTATTTGGAATTTCCCTCGCCCACTCTTACCTAATTTGAATTTGTTGTA	840
Db	829	CGAACCTATTGGGATTTTGAATTTCCCTCGCCCACTCTTACCAATGTTGATTTTGTGGA	888
Qy	841	AGACTCTACTGCAAACCTGTCAACCCCTGCCTAAGGAGAAAAATGGAAGAATTTGCCAG	900
Db	889	GGACTTCACTGTAAACAGCCAAACCCCTGCCTAAGG---AAATGGAAGATTTGTGCAG	945
Qy	901	AGCTCTGATGAAGACGG---TGTTGTGTTTTCTCTGGAGTCAGCTGTGCAAAACCTTACA	957
Db	946	AGCTCTGGAGAAAATGTTATTGTGGTGTCTCTCTGGGTCGATGATCAGTAACATGTCA	1005
Qy	958	GAGAAAAAGCTGATCTTATCACTTCGCGCCCTGGCTCAGATTCCACAAAAAGTCATGAAG	1017
Db	1006	GAGAAAGTGCCAACATGATTGCATCAGCCCTTGCCAGATCCACAAAAGGTTCTATGG	1065
Qy	1018	-----TTCGGAAGGAAACCAATACCTTAAGATCCAATACTCAGTGGCATAGGTGGATC	1071
Db	1066	AGATTTGATGGCAAGAAGCCAAATACCTTAGGTTCCAATACTCGACTGTATAAGTGGTTA	1125
Qy	1072	CCACAGAATGAATGTCTTATCTAGATCATCCCAACCAAGCCTTTATAACTTATGGT	1131

XX
 PN WO200006776-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 22-JUL-1999; 99WO-US016675.
 XX
 PR 28-JUL-1998; 98US-0094391P.
 XX
 PA (AXYS-) AXYS PHARM INC.
 XX
 PI Galvin M, Miller A, Penny L, Riedy M;
 XX
 DR WPI; 2000-195321/17.
 DR P-PSDB; AAY78934.
 XX
 PT Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
 PT genotyping individuals to predict rate of metabolism of substrates and
 PT for identifying potential drug interactions.
 XX
 PS Disclosure; Page 41-44; 72pp; English.
 XX
 CC This sequence represents the human UDP-glucuronosyltransferase 2B7
 CC (UGT2B7) gene. UDP-glucuronosyltransferase (UGTs) are a family of enzymes
 CC that catalyse the glucuronic acid conjugation of a wide range of
 CC endogenous and exogenous substrates. The UGT2B gene subfamily encode
 CC steroid metabolizing isoforms in the liver. Alteration of the expression
 CC or function of UGTs may effect drug metabolism. The invention relates to
 CC non-chromosomal nucleic acid molecules, which comprise human UGT2B
 CC sequence polymorphisms (see AAZ95051-Z95110). Probes which detect the
 CC UGT2B locus polymorphisms can be used to detect altered UGT2B metabolism
 CC of a substrate in an individual. The nucleic acid molecules comprising a
 CC human UGT2B sequence polymorphism can be used in screening assays for
 CC genotyping individuals, also to predict their rate of metabolism of UGT2B
 CC substrate, potential drug-drug interactions and adverse side effects. The
 CC polymorphisms can be used as single nucleotide polymorphisms (SNPs) for
 CC detecting genetic linkage related to phenotypic variation in activity or
 CC expression of UGT2B protein. The polymorphism containing nucleic acid
 CC molecules may also be used for generating genetically modified non-human
 CC animals and for obtaining site specific gene modification in cell lines
 XX
 SQ Sequence 1854 BP; 572 A; 338 C; 392 G; 552 T; 0 U; 0 Other;

Query Match 60.2%; Score 966.6; DB 3; Length 1854;
 Best Local Similarity 78.1%; Pred. No. 4.2e-241;
 Matches 1261; Conservative 0; Mismatches 324; Indels 30; Gaps 7;

Qy	1	ATGGCTATGAAATGGACTTCAGTCCTTCTGTTGATACAGCTGAGCTATTACTCTAGCTCT	60
Db	15	ATGTCTGTGAAATGGACTTCAGTAATTTTGCTAATACTAGCTTTTGCTTTAGCTCT	74
Qy	61	GGGAGTTGTGGAAATGTGCCGCTGTGGCCCATGGAATATAGTCCTTGGATGAATATAAAG	120
Db	75	GGGAATTGTGGAAGGTGCTGGTGTGGCAGCAGAATACAGCCATTGGATGAATATAAAG	134
Qy	121	ACAATCCTGGATAAACTTATGCAGATAAGTCATGAGGTGACTGTTCTAACATTGTCAGCT	180
Db	135	ACAATCCTGGATGAGCTTATTCAGAGAGGTCATGAGGTGACTGTACTGGCATCTTCAGCT	194
Qy	181	TCCATTCTTGTTGATCCCAACATAACATCTGTTACTAAATTTGAGGTTTATTCTATATCT	240
Db	195	TCCATTCTTTTGTATCCCAACAACTCATCCGCTCTTAAATTTGAAATTTATCCACATCT	254
Qy	241	GTAATTAAAGATGATTTTGCAGGGTTTTTTTTCACACAACAGATTACTAAATGGATACAT	300
Db	255	TTAACTAAAGATGAGTT---GGAGAAATTCATCATGCAACAGATTAAGAGATGGTCA---	308
Qy	301	GATCTTCCAAAACATATATTTTGGTTTAAATGTGTTCCCTTCAAGAATATTCTTTGGGAA	360
Db	309	GACCTTCCAAAAGATACATTTTGGTTATATTTTTCACAAGTACAGGAAATCATGTCAATA	368

QY 361 TATTCTGGTTATACTGAGAAGTTCTTTAAAGATGTAGTTTTGAACAAGAACTTATGACA 420
 Db 369 TTTGGTGACATAACTAGAAAGTTCTGTAAAGATGTAGTTTCAAATAAGAAATTTATGAAA 428

QY 421 AACCTACAAGAATCAAGGTCTGATGTCGTTTCATGCAAAATGCCATTGGTCCCTTTGGAGAG 480
 Db 429 AAAGTACAAGAGTCAAGATTTGACGTCATTTTGCAGATGCTATTTTCCCTGTAGTGAG 488

QY 481 CTGCTGGCTGAGCTATTAAAAATATCCTTTGTGTACAGTCTCCACTTCTCTCCTGGCTAC 540
 Db 489 CTGCTGGCTGAGCTATTTAACATACCCCTTTGTGTACAGTCTCAGCTTCTCTCCTGGCTAC 548

QY 541 ACATTTGAGAAATACAGTGGAGGATTTCTACTTCCACCTTCCTATGGAGCTGTTATTCTG 600
 Db 549 ACTTTTGAAAAGCATAGTGGAGGATTTATTTCCCTCCTTCCTACGTACCTGTTGTATG 608

QY 601 TCAGAAATTAAGTGGTTCGATGACATTCATGGAGACAGTAAGAAATATATATATGTGTTT 660
 Db 609 TCAGAAATTAAGTGCATCAATGACTTTTCATGGAGAGGGTAAAAATATGATCTATGTGCTT 668

QY 661 TATTTTGACTTTTGGTTCCAAACATTTTGATATGAAGAAGGGAGACCAGTTTACAGTGAA 720
 Db 669 TACTTTGACTTTTGGTTGAAATATTTGACATGAAGAAGTGGATCAGTTTATAGTGAA 728

QY 721 GTTCTAGGTAAGTCATGTTTTTTATCTGAGATAATGGGAAAAGCTGAAATGTGGCTCATT 780
 Db 729 GTTCTAGGAAGACCCACTACATTATCTGAGACAATGGGAAAAGCTGACGTATGGCTTATT 788

QY 781 CGAAACTACTGGTATTGGAATTTCTCGCCCACTCTTACCTAATTTGAATTTGTTGTA 840
 Db 789 CGAAACTCCTGGAATTTTCAGTTTCCATATCCACTCTTACCAATGTTGATTTTGTGGA 848

QY 841 AGACTCTACTGCAAACCTGTCAACCCCTGCCTAAGGAGAAAATGGAAGAATTTGCCAG 900
 Db 849 GGACTCCACTGCAAACCTGCCAAACCCCTGCCTAAGG---AAATGGAAGACTTTGTACAG 905

QY 901 AGCTCTGATGAAGACGGTGT---GTGTTTTCTCTGGAGTCAGCTGTGCAAAACCTTACA 957
 Db 906 AGCTCTGAGAAAATGGTGTGTGGTGT---TCTCTGGGGTCAATGGTCAGTAACATGACA 965

QY 958 GAAGAAAAGCTGATCTTATCACTTCGGCCCTGGCTCAGATTCACAAAAAGTCATGAAG 1017
 Db 966 GAAGAAAAGGCCAACGTAATTGCATCAGCCCTGGCCAGATCCACAAAAGGTTCTGTGG 1025

QY 1018 -----TTCGGAAGGAAACCAATACCTTAAGATCCAATACTCAGTGGCATAGGTGGATC 1071
 Db 1026 AGATTTGATGGGAATAAACAGATACCTTAGGTCTCAATACTCGGCTCTACAAGTGGA 1085

QY 1072 CCACAGAATGAATGTCTTATCCTAGATCATCCCCAAACCAAGCCTTTATAACTTATGGT 1131
 Db 1086 CCCAGAATGA-----CCTTCTAGGTATCCAAAGACAGAGCTTTATAACTCATGGT 1139

QY 1132 GGAACAAATAGCATCTATGAGATGATCTACCGTGGAGTCCCTTCCATGGGCATTCCCTTG 1191
 Db 1140 GGAGCCAATGGCATCTACGAGGCAATCTACCATGGGATCCCTATGGTGGGGATTCCATTG 1199

QY 1192 TTTGCGGACCAACATGATAACATTGCTCACATGAAGGCCAAGGGAGCAGCTGTTATATTG 1251
 Db 1200 TTTGCCGATCAACCTGATAACATTGCTCACATGAAGGCCAGGGAGCAGCTGTTAGAGTG 1259

QY 1252 GACTTGAGCACAAAGTCAAGTACAGATTTGCTCGATATATCTGTGTTTCGTATCTTTATTT 1311
 Db 1260 GACTTCAACACAATGTCGAGTACAGACTTGCTGAATGCATTGAAGAGAGTAATTAATGAT 1319

QY 1312 TTATCCTTCAGATATAAAGAGAGTGTTATGAAATTATCAAGAATTCAACATGATCAACCA 1371
 Db 1320 CCTTC-----ATATAAGAGAATGTTATGAAATTATCAAGAATTCAACATGATCAACCA 1373

QY 1372 GTGAAGCCCCTGGATCGAGCAGTCTTCTGGATTGAATTTGTCATGCGCCACAAAGGAGCC 1431
 Db 1374 GTGAAGCCCCTGGATCGAGCAGTCTTCTGGATTGAATTTGTCATGCGCCACAAAGGAGCT 1433

Qy 1432 AAACACCTTCGAGTTGCAGCCCGTGACCTCACCTGGTTCCAGTACCACTCTTTGGATGTG 1491
 |||||
 Db 1434 AAACACCTTCGGGTTGCAGCCACGACCTCACCTGGTTCCAGTACCACTCTTTGGATGTG 1493
 |||||
 Qy 1492 ATTGGGTTTCTGCTGGCCTGTGTGGCAACTGTGACATTTATCATCACAAAGTGTGTCTG 1551
 |||||
 Db 1494 ATTGGGTTCTGCTGGTCTGTGTGGCAACTGTGATATTTATCGTCACAAAATGTTGTCTG 1553
 |||||
 Qy 1552 TTTTGTCTTCTGGAAGTTTACTAGAAAAGTGAAGAAGGAAAAAAGGGATTAGTTAT 1606
 |||||
 Db 1554 TTTTGTCTTCTGGAAGTTTGTCTAGAAAAGCAAAGAAGGGAAAAAATGATTAGTTAT 1608
 |||||

RESULT 11

AAV15900

ID AAV15900 standard; cDNA; 2107 BP.

XX

AC AAV15900;

XX

DT 26-MAY-1998 (first entry)

XX

DE Uridine diphospho-glucuronosyltransferase 2B17 (UGT2B17) encoding cDNA.

XX

KW Uridine diphospho-glucuronosyltransferase 2B17; UGT2B17; catalyse;
 KW androsterone; androsterone-glucuronic acid; androgen; enzyme; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	5'UTR	1. .51
FT		/*tag= a
FT	CDS	52. .1644
FT		/*tag= b
FT		/product= "UGTB17 enzyme"
FT	3'UTR	1645. .2107
FT		/*tag= c

XX

PN WO9744466-A1.

XX

PD 27-NOV-1997.

XX

PF 16-MAY-1997; 97WO-CA000328.

XX

PR 17-MAY-1996; 96US-00649319.

XX

PA (ENDO-) ENDORECHERCHE INC.

XX

PI Belanger A, Hum DW, Beaulieu M, Levesque E;

XX

DR WPI; 1998-018520/02.

DR

P-PSDB; AAW47126.

XX

PT DNA encoding uridine di:phospho:glucuronosyl:transferase 2B17 - which
 PT catalyses conversion of androsterone to androsterone-glucuronic acid.

XX

PS Claim 15; Page 4-6; 53pp; English.

XX

CC This cDNA encodes an enzyme uridine di-phosphoglucuronosyltransferase
 CC 2B17 (UGT2B17). This novel enzyme catalyses the conversion of
 CC androsterone to androsterone-glucuronic acid. The UGT2B17 can be used to
 CC detect anti-UGT2B17 antibodies. The antibody can be used to detect a
 CC localised concentration of UGT2B17 or an alteration in androgen activity.
 CC The UGT2B17 can also be used to alter the concentration of an androgenic
 CC compound in a tissue, specifically dihydrotestosterone. An isolated
 CC nucleotide sequence comprising at least 30 consecutive nucleotides from
 CC the coding region of the 2107 base pair sequence, or its complement can
 CC be used to block the synthesis of UGT2B17, e.g. an expression disrupting
 CC sense or antisense fragment, or as a probe for a UGT2B17 coding sequence

SQ Sequence 2107 BP; 657 A; 382 C; 424 G; 644 T; 0 U; 0 Other;

Query Match 59.7%; Score 958.6; DB 2; Length 2107;
Best Local Similarity 77.3%; Pred. No. 5.4e-239;
Matches 1249; Conservative 0; Mismatches 339; Indels 27; Gaps 6;

QY	1	ATGGCTATGATAAAATGGACTTTCAGTCTCTTCTGTTGATACAGCTAGCTATTACTCTAGCTCT	60
Db	52	ATGTCTCTGAAATGGATGTCAGTCTTTCTGCTGATGCAGCTCAGTTGTTACTTTAGCTCT	111
QY	61	GGGAGTTGTGGAAATGTGCCGCTGTGGCCCATGGAATATAGTCCTTGGATGAATATAAAG	120
Db	112	GGGAGTTGTGGAAAGGTGCTGGTGTGGCCACAGAATACAGCCATTGGATAAATATGAAG	171
QY	121	ACAATCCTGGATAAACTTATGCAGATAAGTCATGAGGTGACTGTTTCAACATTGTGAGCT	180
Db	172	ACAATCCTGGAAGAGCTTGTTTCAGAGGGGTCATGAGGTGATTGTGTTGACATCTTCGGCT	231
QY	181	TCCATTCTTGTTGATCCCAACATAACATCTGTTACTAAATTTGAGGTTTATTCTATATCT	240
Db	232	TCTATTCTTGTCATGCCAGTAAATCATCTGCTATTAAATTAGAAGTTTATCCTACATCT	291
QY	241	GTAATTAAGATGATTTTGCAGGGTTTTTTTTTCACACAACAGATTACTAAATGGATACAT	300
Db	292	TTAACTAAAAATGATTGGAAGATTTTTTTATGA---AAATGTTTCGATAGATGGACATAT	348
QY	301	GATCTTCCAAAAACATATATTTTGGTTTAAATGTGTCCCTTCAAGAATATTCTTTGGGAA	360
Db	349	AGTATTTCAAAAAATACATTTTGGTCATATTTTTCACAACTACAAGAATTGTGTTGGGAA	408
QY	361	TATTCTGGTTTATACTGAGAAGTTCTTTAAAGATGTAGTTTGAACAAGAACTTATGACA	420
Db	409	TATTCTGACTATAATATAAAGCTCTGTGAAGATGCAGTTTGAACAAGAACTTATGAGA	468
QY	421	AACCTACAAGAATCAAGGTCTGATGTCGTTTCATGCAAATGCCATTGGTCCCTTTGGAGAG	480
Db	469	AAACTACAAGAGTCAAAATTTGATGTCTTCTGGCAGATGCCGTTAATCCCTGTGGTGAG	528
QY	481	CTGCTGGCTGAGCTATTAATAAATATCCTTTGTGTACAGTCTCCACTTCTCTCCTGGCTAC	540
Db	529	CTGCTGGCTGAACACTACTTAACATACCCTTCTGTACAGTCTCCGCTTCTCTGTGGCTAC	588
QY	541	ACATTGAGAAATACAGTGGAGGATTTCTACTTCCACCTTCTATGGAGCTGTTATTCTG	600
Db	589	ACAGTTGAGAAGAATGGTGGAGGATTTCTGTTCCCTCCTTCTATGTACCTGTTGTTATG	648
QY	601	TCAGAATTAAGTGGTTCGATGACATTCATGGAGACAGTAAGAAATATTATATATGTGTTT	660
Db	649	TCAGAATTAAGTGATCAAATGATTTTCATGGAGAGGATAAAAAATATGATATATATGCTT	708
QY	661	TATTTTGACTTTTGGTTCCAACATTTGATATGAAGAAGGGAGACCAGTTTACAGTGAA	720
Db	709	TATTTTGACTTTTGGTTTCAAGCATATGATCTGAAGAAGTGGGACCAGTTTATAGTGAA	768
QY	721	GTTCTAGGTAAGTCATGTTTTTTATCTGAGATAATGGGAAAAGCTGAAATGTGGCTCATT	780
Db	769	GTTCTAGGAAGACCCACTACATTATTTGAGACAATGGGAAAAGCTGAAATGTGGCTCATT	828
QY	781	CGAAACTACTGGTATTTGGAATTTCTCGCCCACTCTTACCTAATTTTGAATTTGTTGTA	840
Db	829	CGAACCTATTGGGATTTTGAATTTCTCGCCCACTCTTACCAAATGTTGATTTTGTGGA	888
QY	841	AGACTCTACTGCAAACCTGTCAACCCCTGCCTAAGGAGAAAAATGGAAGATTTGCCAG	900
Db	889	GGACTTCACTGTAAACAGCCAAACCCTTGCTAAGG---AAATGGAAGAGTTGTGCAG	945
QY	901	AGCTCTGATGAAGACGG---TGTTGTGTTTTCTCTGGAGTCAGCTGTGCAAAACCTTACA	957
Db	946	AGCTCTGGAGAAAATGGTATTGTGGTGTTTTCTCTGGGGTCGATGATCAGTAACATGTCA	1005

Result No.	Score	Query Match	Length	DB	ID	Description
1	798.4	49.7	2079	14	CD013998	CD013998 90117389
2	784.2	48.8	1946	11	AK050435	AK050435 Mus muscu
3	777.8	48.4	2573	11	AK004971	AK004971 Mus muscu

4	713	44.4	1892	11	AK034801
5	710.4	44.2	1896	11	AK002736
6	695.4	43.3	2895	11	AK083294
7	689.8	43.0	1642	14	CD013996
8	600.8	37.4	1716	14	CD013997
9	554.6	34.5	1783	11	AK050327
10	549.8	34.2	2099	11	AK008601
11	525	32.7	983	13	BX444042
12	520.8	32.4	1023	14	CD013994
13	520.8	32.4	1336	14	CD013995
14	518.6	32.3	2575	11	BC048920
c 15	442	27.5	823	10	BF689099

AK034801	Mus muscu
AK002736	Mus muscu
AK083294	Mus muscu
CD013996	90117309
CD013997	90117357
AK050327	Mus muscu
AK008601	Mus muscu
BX444042	BX444042
CD013994	90130114
CD013995	90130122
BC048920	Mus muscu
BF689099	602185172